Journal of Symbolic Computation 44 (2009) 1017–1036

Contents lists available at ScienceDirect

Journal of Symbolic Computation

journal homepage: www.elsevier.com/locate/jsc

On the Moser- and super-reduction algorithms of systems of linear differential equations and their complexity

Moulay A. Barkatou [a], Eckhard Pflügel [b]

[a] Université de Limoges, CNRS, XLIM (UMR 6172), Département Mathématiques Informatique, 123, Av. A. Thomas, 87060 Limoges, France

[b] Faculty of CISM, Kingston University, Penrhyn Road, Kingston upon Thames, Surrey KT1 2EE, United Kingdom

ARTICLE INFO

Article history:
Received 6 September 2008
Accepted 15 January 2009
Available online 23 February 2009

Keywords:
Computer algebra
Local analysis of linear differential systems
Moser-reduction
Super-reduction
Singularities

ABSTRACT

The notion of irreducible forms of systems of linear differential equations with formal power series coefficients as defined by Moser [Moser, J., 1960. The order of a singularity in Fuchs' theory. Math. Z. 379–398] and its generalisation, the super-irreducible forms introduced in Hilali and Wazner [Hilali, A., Wazner, A., 1987. Formes super-irréductibles des systèmes différentiels linéaires. Numer. Math. 50, 429–449], are important concepts in the context of the symbolic resolution of systems of linear differential equations [Barkatou, M., 1997. An algorithm to compute the exponential part of a formal fundamental matrix solution of a linear differential system. Journal of App. Alg. in Eng. Comm. and Comp. 8 (1), 1–23; Pflügel, E., 1998. Résolution symbolique des systèmes différentiels linéaires. Ph.D. Thesis, LMC-IMAG; Pflügel, E., 2000. Effective formal reduction of linear differential systems. Appl. Alg. Eng. Comm. Comp., 10 (2) 153–187]. In this paper, we reduce the task of computing a super-irreducible form to that of computing one or several Moser-irreducible forms, using a block-reduction algorithm. This algorithm works on the system directly without converting it to more general types of systems as needed in our previous paper [Barkatou, M., Pflügel, E., 2007. Computing super-irreducible forms of systems of linear differential equations via Moser-reduction: A new approach. In: Proceedings of ISSAC'07. ACM Press, Waterloo, Canada, pp. 1–8]. We perform a cost analysis of our algorithm in order to give the complexity of the super-reduction in terms of the dimension and the Poincaré-rank of the input system. We compare our method with previous algorithms and show that, for systems of big size, the direct block-reduction method is more efficient.

E-mail addresses: Moulay.Barkatou@xlim.fr (M.A. Barkatou), E.Pfluegel@kingston.ac.uk (E. Pflügel).

0747-7171/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsc.2009.01.002

1. Introduction

Let F be a subfield of the field $\mathbb{C}$ of complex numbers (for example $\mathrm{F} = \mathbb{Q}$ or an algebraic extension of $\mathbb{Q}$) and $\overline{\mathrm{F}}$ an algebraic closure of F. Denote by $\mathcal{O} = \mathrm{F}[[x]]$ the ring of formal power series in x with coefficients in F. Let $\mathrm{K} = \mathrm{F}[[x]][x^{-1}]$ be the quotient field of $\mathcal{O}$. If $a \in \mathrm{K}$, we denote the order in x of a by $\mathrm{v}(a)$, $(\mathrm{v}(0) = +\infty)$. The quantity $\mathrm{v}(a)$ is also called the *valuation* of a. If $A = (a_{i,j})$ is a matrix with coefficient in K, we define the valuation of A by

$$\mathrm{v}(A) = \min_{i,j} (\mathrm{v}(a_{i,j})).$$

Throughout this paper we let ϑ denote the derivation $x\frac{\mathrm{d}}{\mathrm{d}x}$ of K and consider a formal first-order linear differential system of the form

$$\vartheta y = A(x)y, \tag{1}$$

where y is a vector with $n \geq 2$ components, and $A(x)$ is an $n \times n$ matrix whose coefficients are in K and write

$$A(x) = x^{-\mathrm{v}(A)} \sum_{j=0}^{\infty} A_j x^j$$

where the $A_j \in \mathrm{F}^{n\times n}$ are constant $n \times n$ matrices with $A_0 \neq 0$. The non-negative integer

$$r = r(A) := \max(0, -\mathrm{v}(A))$$

will be called the *Poincaré-rank* of (1).

The change of variable $y = T(x)z$ where the *transformation matrix* $T \in \mathrm{GL}(n, \mathrm{K})$ transforms (1) into a new system

$$\vartheta z = B(x)z \tag{2}$$

where

$$B = T[A] := T^{-1}AT - T^{-1}\vartheta T. \tag{3}$$

We call the systems (1) and (2) (the matrices A and B respectively) *equivalent*.

Given a differential system (1) with Poincaré-rank $r > 0$, one is interested in finding an equivalent system whose Poincaré-rank is minimal. This motivates the notion of reducibility of a system as defined in Moser (1960): let

$$m(A) = r + \frac{\operatorname{rank} A_0}{n}$$

and

$$\mu(A) = \min_{T\in\mathcal{T}_n} \left\{ m\left(T[A]\right) \right\}$$

where $\mathcal{T}_n = \mathrm{GL}(n, K)$ denotes the set of transformation matrices.

Definition 1.1. The system (1) (the matrix A respectively) is called *Moser-reducible* if $m(A) > \mu(A)$, otherwise it is said to be *Moser-irreducible*.

Having efficient algorithms for computing Moser-irreducible forms of this type of system is important and has given rise to a range of applications in computer algebra: the problem of *classification of singularities* – which was Moser's initial motivation for his work – is solved by computing an equivalent Moser-irreducible system. If the singularity is regular, this will result in a system with simple pole. In this situation, classical algorithms can then be used in order to compute formal solutions in the form of formal power series, mixed with logarithmic terms. In particular, this will reveal the *indicial equation* of the system, whose integer roots give valuable information about the structure of polynomial and rational solutions of systems with coefficients in $\mathrm{F}(x)$.

In Barkatou (1999), the first author showed that in the general case, when dealing with an irregular singularity, the so-called *super-irreducible* forms of linear differential systems are useful for computing the indicial equation, and aid in finding efficient algorithms for computing formal power series solutions, polynomial or rational solutions.

Moser-irreducible forms have been used by the authors of the present paper in Barkatou (1997) and Pflügel (2000) in order to compute exponential parts of formal solutions in the case of an irregular singularity.

In this paper, we make a new contribution by extending our previous paper (Barkatou and Pflügel, 2007). We analyze the algorithm we have given therein for computing Moser-irreducible forms and give its complexity. We develop a new version of the block-reduction algorithm presented in the same previous paper, that works directly on the system's coefficients. We use this to reduce the task of computing super-irreducible forms to that of computing one or several Moser-irreducible forms. Finally, in Section 5 we give a time comparison of our method in Barkatou and Pflügel (2007) (Method A) and the method of the present paper (Method B). The obtained data suggests that Method B is more efficient for systems whose dimension is large.

The algorithms discussed and developed in this paper manipulate matrices with power series coefficients. Normally, computations would have to be done on an infinite number of coefficients in order to obtain exact results. However, in practice only a finite number of coefficients is necessary for computing Moser- and super-irreducible forms and we will work with truncated power series. We will denote by ν the number of coefficients that are taken into account. From a theoretical point of view, one can take $\nu = nr$ (this is enough if one wants to compute the exponential parts, see for example (Babitt and Varadarajan, 1983; Barkatou, 1997). In our implementation, this problem is solved using lazy evaluation.

Note that the cyclic vector method (Barkatou, 1993; Churchill and Kovacic, 2002) is an alternative approach for dealing with the symbolic resolution of linear differential systems in general, and the problem of computing local invariants in particular. From a practical point of view, various authors have pointed out in the past that direct methods tend to perform better than the cyclic vector approach, especially when n is big — see Pflügel (1998) for a comparison and some timings, or Cluzeau (2003) for a discussion of computing rational solutions. From a theoretical point of view the complexity of the cyclic vector method is $O(n^5)$ operations in $\mathrm{K} = \mathrm{F}((x))$. As shown in Cluzeau (2003), the degree of the coefficients of the resulting scalar nth-order linear differential equation is bounded by $O(n^3)$ if we take into account that one needs the first $\nu = rn$ coefficients of A, which suggests that the complexity of the cyclic vector method is worse than the complexity of our algorithm.

For example, consider the matrix

$$A = \begin{bmatrix} 2\,x^{-6} & -3 & \dfrac{-1+3\,x^3}{x^8} & -2\,x^{-2} \\ 0 & -\dfrac{4+x^2}{x^4} & -2\,x^{-6} & -2\,\dfrac{x^4-1}{x^9} \\ \dfrac{5\,x^2+2}{x^7} & -\dfrac{-5+2\,x}{x^7} & -x^{-2} & -6 \\ -4\,x^{-1} & -3 & 6 & -6\,x^{-1} \end{bmatrix}.$$

An attempt to convert the third symmetric power matrix of A, a matrix of dimension 20, using the cyclic vector method had to be interrupted unsuccessfully whereas the direct methods for the super-reduction terminate after just above 60 s.

Notation

By $\mathcal{M}_n(\mathcal{O})$, $\mathcal{M}_n(\mathrm{K})$ we denote the ring of $n \times n$ matrices whose elements lie in $\mathcal{O}$, K respectively. We write $\mathrm{GL}(n, \mathrm{K})$ for the group of invertible matrices in $\mathcal{M}_n(\mathrm{K})$.

By 0_s we denote the zero square matrix of size s and by I_s the identity matrix of dimension s. By $\mathrm{diag}(a, b, c, \ldots)$ we denote the square-diagonal (block-diagonal respectively) matrix whose diagonal elements are $a, b, c, \ldots$.

We also denote by ω the exponent for the complexity of matrix multiplication (Von zur Gathen and Gerhard, 2003) such that two matrices in $\mathcal{M}_n(\mathrm{F})$ can be multiplied using $\mathrm{O}(n^{\omega})$ operations in F. For the standard matrix multiplication, one has $\omega = 3$.

2. The Moser-reduction

2.1. A reduction criterion

In order to design effective reduction algorithms, one needs a constructive criterion to decide whether or not a given system is Moser-reducible. Following Moser, we define the *associated polynomial*

$$\theta(A, \lambda) = x^{\operatorname{rank} A_0} \det\left(x^{-1}A_0 + A_1 - \lambda I\right)\Big|_{x=0}.$$

It has been shown in Moser (1960) that a matrix $A \in \mathcal{M}_n(K)$ with Poincaré-rank $r \geq 1$ is Moser-reducible if and only if $\theta(A, \lambda) \equiv 0$ (note that our definition of the associated polynomial differs from that in Moser (1960) by replacing λ with $-\lambda$).

In what follows, we assume that A_0 is nilpotent. This assumption is not very restrictive, since it is in fact a necessary condition for the existence of a transformation which lowers the Poincaré-rank r (Moser, 1960). The case where A_0 has several eigenvalues can be reduced to the nilpotent case by using a constant similarity transformation which puts A_0 into block-diagonal form

$$A_0 = \begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix} \tag{4}$$

where M is invertible and N nilpotent. Since M and N have no common eigenvalue, the well-known *Splitting Lemma* (Wasow, 1967) (see the appendix) states that there exists a formal transformation matrix

$$T(x) = \sum_{j=0}^{\infty} T_j x^j \quad (T_0 = I)$$

such that the change of variable $y = Tz$ transforms the system (1) into a new system

$$\vartheta z = B(x)z$$

where

$$B = \begin{pmatrix} B^{11} & 0 \\ 0 & B^{22} \end{pmatrix}$$

is of same Poincaré-rank r and block-diagonal with the same block partition as in A_0. It can be shown that the system corresponding to the first block is Moser-irreducible. The second system has a nilpotent leading matrix, we have $\operatorname{rank} A_0 = \dim M + \operatorname{rank} N$ and applying the reduction algorithm to the second system minimises $\operatorname{rank} A_0$. Note that in practice, the block-diagonalisation only needs to be computed up to a small order, using lazy evaluation. In the appendix we will review this algorithm and give its complexity.

Throughout this paper, we suppose that $r \geq 1$ and that the nilpotent matrix A_0 is in canonical Jordan normal form. This can always be achieved by using a constant transformation matrix over F, and there are efficient algorithms for this task (Giesbrecht and Storjohann, 2002). Thus,

$$A_0 = \operatorname{diag}(J, 0_s) \tag{5}$$

where J has d Jordan blocks of dimension $n_i \geq 2$ (with $n_1 \geq \cdots \geq n_d > n_{d+1} = \cdots = n_{d+s} = 1$) and define for $i = 1, \ldots, d + s$ the positive integers l_i (c_i respectively) as the position of the ith zero row (column respectively) of A_0. We also define similarly the positive integers l_i^* and c_i^*, for $i = 1, \ldots, n - d - s$, as positions of the non-zero rows and columns of A_0. Furthermore, denote by $A_L \in \mathcal{M}_{d+s}(\mathcal{O})$ the submatrix of $x^r A$ whose entries are given by the elements of positions (l_i, c_j) of $x^r A$.

Note that we have

$$A_L = A_{L,1}x + O(x^2) \text{ with } A_{L,1} \in M_{d+s}(\mathrm{F}). \tag{6}$$

The definition of the *L-Matrix* $L(A,\lambda) \in \mathcal{M}_{d+s}(\mathrm{F}[\lambda])$ we introduced in Barkatou and Pflügel (2007) can be stated as

$$L(A,\lambda) = A_{L,1} - \mathrm{diag}(0_d, \lambda I_s). \tag{7}$$

One then has the following

Proposition 2.1 (*Barkatou and Pflügel, 2007*). *The system* (1) *is Moser-reducible if and only if* $\det L(A,\lambda) \equiv 0$.

2.2. Moser-reduction revisited

In this section, we present the algorithm for the Moser-reduction that we have first developed in Barkatou and Pflügel (2007). We follow closely the presentation of the initial paper. New results in this section are Lemma 2.4 and Proposition 2.2 which will be used for the complexity analysis of our super-reduction algorithm in Section 4.3.

Our algorithm works by first bringing the given system into a convenient form using a constant transformation, and then carrying out a reduction step using a diagonal transformation, which decreases the quantity $m(A)$.

One can show (Barkatou and Pflügel, 2007) that this step produces a maximal reduction of the rank of A_0 in the sense of Dietrich (1978).

Lemma 2.1. *Suppose that* A *is Moser-reducible. There exists a constant transformation* C *such that* $\tilde{A} := C[A]$ *with* $\tilde{A}_0 = A_0$ *and*

$$L(\tilde{A},\lambda) = \begin{pmatrix} L^{11} & L^{12} & 0 \\ L^{21} & L^{22}-\lambda & 0 \\ L^{31} & L^{32} & L^{33}-\lambda \end{pmatrix} \tag{8}$$

where L^{11}, L^{22} *and* L^{33} *are square matrices of dimension* d, $s-q$ *and* q *with* $0 \le q \le s$, *furthermore*

$$\mathrm{rank}\begin{pmatrix} L^{11} \\ L^{21} \end{pmatrix} + s - q = \mathrm{rank}\begin{pmatrix} L^{11} & L^{12} \\ L^{21} & L^{22} \end{pmatrix} \tag{9}$$

and L^{33} *is lower triangular with zero diagonal.*

Proof. We shall reason inductively. Let $q = 0$ and $L(A,0)$ be partitioned

$$L(A,0) = \begin{pmatrix} L^{11} & L^{12} \\ L^{21} & L^{22} \end{pmatrix}, \tag{10}$$

so that L^{11} and L^{22} are square matrices of order d and s respectively. Let E (respectively F) be the vector space spanned by the first d (respectively the last s) columns of $L(A,0)$. Since A is Moser-reducible, the matrix $L(A,0)$ is singular. Hence one has

$$\dim(E+F) = \mathrm{rank}\, L(A,0) < d+s.$$

If $\dim E + s = \mathrm{rank}\, L(A,0)$ then take $C = I_n$. Otherwise we have $\dim E + s > \mathrm{rank}\, L(A,0)$. Using the fact that $\dim(E+F) + \dim(E\cap F) = \dim E + \dim F$, we see that $\dim F < s$ or $\dim(E \cap F) > 0$. This implies that the matrix $L(A,0)$ must have at least one column with index $d < i \le d+s$ which is a linear combination of columns with index $\ne i$. By using a constant transformation which swaps rows and columns we can achieve $i = d+s$. Note that this transformation preserves the Jordan structure of A_0.

It is easily verified that we can now eliminate the last column of $L(A,0)$ through a constant transformation using row and column eliminations on A. To achieve this, let ${}^t v = (v_1,\ldots,v_{d+s}) \in \ker L(A,0)$ with $v_{d+s} = 1$ and define

$${}^t u = (v_1, \overbrace{0,\ldots,0}^{n_1-1}, \ldots, v_d, \overbrace{0,\ldots,0}^{n_d-1}, v_{d+1}, \ldots, v_{d+s-1}).$$

The constant transformation is then

$$P = \begin{pmatrix} I_{n-1} & u \\ 0 & 1 \end{pmatrix}.$$

Let $\widetilde{A}$ denote the matrix of the resulting system. Then $\widetilde{A}_0 = A_0$ and $\widetilde{A}_1$ has its last column zero. Thus $L(\widetilde{A}, \lambda)$ has the form (8) with $q = 1$ and $L^{33} = 0_1$, hence trivially L^{33} is upper triangular with zero diagonal. If the condition (9) is not satisfied, then developing the determinant of $L(\widetilde{A}, \lambda)$ shows that the submatrix $\begin{pmatrix} L^{11} & L^{12} \\ L^{21} & L^{22} \end{pmatrix}$ must be singular. One can hence repeat the same process and increase q by 1. After a finite number of iterations of this process we obtain an equivalent matrix $\widetilde{A}$ for which the condition (9) occurs or $q = s$. But in the latter case one has $\det L^{11} = 0$, and (9) holds trivially. □

Lemma 2.2. *Suppose that A is Moser-reducible and $L(A, \lambda)$ has the structure as in (8), with (9) satisfied. Then*

$$\operatorname{rank} \begin{pmatrix} L^{11} \\ L^{21} \end{pmatrix} < d.$$

Proof. Since A is Moser-reducible, we have

$$0 = \theta(A, \lambda) = \pm\lambda^q \det \tilde{L}(\lambda)$$

where

$$\tilde{L}(\lambda) = \begin{pmatrix} L^{11} & L^{12} \\ L^{21} & L^{22} - \lambda \end{pmatrix}$$

and hence $\operatorname{rank} \tilde{L}(0) < d + s - q$. Using assumption (9) of the lemma, one finds

$$\operatorname{rank} \begin{pmatrix} L^{11} \\ L^{21} \end{pmatrix} = \operatorname{rank} \tilde{L}(0) - s + q < d. \quad \square$$

Remark 2.1. The block-triangular form (8) together with condition (9) improves the so-called qtcd-form in Hilali and Wazner (1987) in two aspects: we consider the matrix $L(A, \lambda)$ which is of smaller size than the matrix used in the qtcd-form, and condition (9), inspired by that used in the algorithm in Dietrich (1978), is stronger than the termination criterion used in the qtcd-form.

Lemma 2.3. *Suppose that A is Moser-reducible and $L(A, \lambda)$ has the structure as in (8), with (9) satisfied. Define*

$$S = \operatorname{diag}(I_{n_1-1}, x, \ldots, I_{n_d-1}, x, xI_{s-q}, I_q). \tag{11}$$

Then $m(S[A]) < m(A)$.

Proof. Let $\tilde{A} = S[A]$. Inspection of the effect of the transformation matrix S shows that $\tilde{A} = x^{-r} \sum_{j=0}^{\infty} \tilde{A}_j x^j$ where $\tilde{A}_0$ has the following structure: for each Jordan block J_i ($i = 1, \ldots, d$) in A_0, copy all 1s except for the last. For $i = 1, \ldots, d + s - q$, the row of index l_i is given by

$$(\overbrace{\alpha_{i,1} * \cdots * 0}^{n_1}, \ldots, \overbrace{\alpha_{i,d} * \cdots * 0}^{n_d}\ \overbrace{0 \ldots 0}^{s}).$$

All remaining entries are zero. From this we can see

$$\operatorname{rank} \tilde{A}_0 = \sum_{i=1}^{d} (n_i - 2) + \operatorname{rank} \begin{pmatrix} L^{11} \\ L^{21} \end{pmatrix} < \sum_{i=1}^{d} n_i - d = \operatorname{rank} A_0. \quad \square$$

This yields the following algorithm:

Moser_reduction(A)
Input: $A \in \mathcal{M}_n K$
Output: $T \in \mathcal{T}_n$ such that $T[A]$ is Moser-irreducible

(1) $T := I_n$;
(2) **while** $(r(A) > 0)$ **and** $(\theta(A,\lambda) \equiv 0)$ **do**
  (a) Compute a constant transformation C such that $L(C[A],\lambda)$ has structure (8);
  (b) $A := C[A]$; $T := TC$;
  (c) Compute a transformation S as in (11);
  (d) $A := S[A]$; $T := TS$;
(3) **return** T;

For ease of presentation, we assume in this description of the algorithm that A_0 is nilpotent in each individual iteration. But this is not restrictive — the definition of the L-matrix can easily be extended to the case where A_0 is block-diagonal as in (4) by taking into account only the second block (the matrix N is supposed to be in Jordan normal form), and adding an identity matrix of appropriate size to the diagonal transformation (11).

2.3. Complexity of the Moser-reduction

Lemma 2.4. *Computing a constant transformation C such that $L(C[A],\lambda)$ has structure (8) together with the ν first terms of $C[A]$ costs $\mathrm{O}(\nu n^3)$ multiplications in* F.

Proof. The computation of C requires at most s steps. At step k we need to compute a vector v in the kernel of a square constant matrix of order $s+d-k+1$. This costs $\mathrm{O}((s+d-k+1)^{\omega})$ operations in F (Von zur Gathen and Gerhard, 2003).

Using the fact that $d+1 \le s+d-k+1 \le (s+d) \le n-1$ and $d \ge 1$, we find that

$$\sum_{k=1}^{s}(s+d-k+1)^{\omega} \le s(s+d)^{\omega} \le n^{\omega+1}.$$

Thus computing the matrix C requires at worst $\mathrm{O}(n^{\omega+1})$ operations in F. At each step k we have to form the matrix P above (see the proof of Lemma 2.1) and compute $P[A]$ up to order ν. Computing $P^{-1}A_iP$ costs $2(d+s-k+1)n$ multiplications due to the particular form of P and P^{-1}. Thus computing $P[A]$ up to order ν costs $2\nu(d+s-k+1)n$. Hence computing $C[A]$ up to order ν costs

$$2\nu n\sum_{k=1}^{s}(s+d-k+1) = 2n\nu\left(\frac{1}{2}s(s+2d+1)\right) \le \nu n(n-2)(n+1)$$

since $s+d \le n-d$ and $d \ge 1$ (recall that d (respectively s) is the number of Jordan blocks of size ≥ 2 (respectively of size 1) in A_0).

The total cost for computing C and $C[A]$ (up to order ν) is then $\mathrm{O}(n^{\omega+1}+\nu n^3)$ which is bounded by $\mathrm{O}(\nu n^3)$, since ν is supposed to be bigger than n and $\omega \le 3$. $\square$

Proposition 2.2. *Consider a system (1) with size n and Poincaré-rank r. Then computing a transformation T such that $T[A]$ is Moser-irreducible together with the ν first terms of $T[A]$ costs $\mathrm{O}(\nu r n^4)$ multiplications in* F.

Proof. Computing the overall transformation T by our Moser-reduction algorithm requires, in the worst case, $r(n-1)$ steps (this occurs when the original system is regular singular). By Lemma 2.4, each step costs $\mathrm{O}(\nu n^3)$ operations in F. $\square$

3. A direct block-reduction algorithm

In Barkatou and Pflügel (2007), we have reduced the computation of a super-irreducible system to that of several Moser-irreducible systems of smaller size, using a block-reduction algorithm. This block-reduction method is based on the *Generalised Splitting Lemma* (Pflügel, 2000) and involves rewriting the given system in a different form as discussed in (Barkatou and Pflügel, 2007, Section 3.2). In this section we give a new version of the block-reduction algorithm which works directly on the coefficient matrix A of the system. The idea is to use elementary operations of the same form as in the computation of the Arnold–Wasow form, using the additional property that the system is in *normalised Moser-irreducible form* (see Section 3.2) . We will give the complexity of this direct algorithm and explain its advantages at the end of this paper.

3.1. On the Arnold–Wasow form

Our initial motivation for investigating the Arnold–Wasow form was for theoretical purposes. However, our main result is a new algorithmic application as described in Section 3.3.

We say that the system (1) is in Arnold–Wasow form,[1] if

$$x^r A(x) = A_0 + \begin{pmatrix} A^{1,1}(x) & \cdots & A^{1,d+s}(x) \\ \vdots & & \vdots \\ A^{d+s,1}(x) & \cdots & A^{d+s,d+s}(x) \end{pmatrix}$$

where $A^{i,i}$ $(i = 1, \ldots, d+s)$ are square matrices of dimension n_i without constant term. Furthermore, $A^{i,j}$ have all zero elements except for the last row if $j \geq i$, and except for the first column if $j < i$. The coefficients of A being power series we also say that A is in Arnold–Wasow form up to order h $(h \in \mathbb{N})$ if the matrix

$$x^{-r}(A_0 + A_1 x + \cdots + A_{h-1} x^h)$$

is in Arnold–Wasow form. The Arnold–Wasow form can be computed using transformations $T \in \mathrm{GL}(n, \mathcal{O})$ consisting of a sequence of elementary row and column operations.

For $\alpha \in \mathcal{O}$ and $1 \leq i, j \leq n$ denote by $E_{i,j}(\alpha)$ the n by n "elementary matrix"

$$E_{i,j}(\alpha) = \begin{pmatrix} 1 & & & & \\ & \ddots & & \alpha & \\ & & \ddots & & \\ & & & 1 & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

where the entry at position (i, j) is α. Note that the inverse of $E_{i,j}(\alpha)$ is $E_{i,j}(-\alpha)$. Transforming the given system (1) with $E_{i,j}(\alpha)$ results in a new system whose coefficient matrix $\tilde{A}$ is obtained from A by adding to the jth column the ith column multiplied by α, then subtracting the jth row multiplied by α from the ith row, and adding $-\vartheta(\alpha)$ to the entry in the (i, j) position.

The algorithm computing the Arnold–Wasow form uses a series of elementary operations of the form $E_{ij}(\sigma_{ij} x^h)$ with $\sigma_{ij} \in \mathrm{F}$ where $h \in \mathbb{N}^*$ if A is in Arnold–Wasow form up to order $h - 1$. This order can be increased to h using the following algorithm:

Arnold_Wasow_Lift[2](A, h)

Input: $A = (a_{pq}) \in \mathcal{M}_n(\mathrm{K})$ in Arnold–Wasow form up to order $h - 1$
Output: $T \in \mathcal{T}_n$ that increases the order to h

(1) $T := I_n$;
(2) **For** k **from** 1 **to** d **do**

---

[1] We recall that A_0 is supposed to be nilpotent in Jordan canonical form.

[2] See Section 2.1 for notations. We also set $l_0 = 0$ and $c_{d+s+1} = n + 1$.

(L_1) **For** each row number i **from** $l_{k-1}+1$ **to** $l_k - 1$ **do**
  (ℓ_1) **For** each column number $j \neq i+1$ **from** c_k **to** n **do**
    (a) $E := E_{i+1,j}(-\alpha_{ij}^* x^h)$; // here $a_{ij} = \alpha_{ij}^* x^{-r+h} + O(x^{-r+h+1})$ with $\alpha_{ij} \in$ F
    (b) $A := E[A]$; // This eliminates the term $\alpha_{ij}^* x^{-r+h}$ in a_{ij}
    (c) $T := TE$;
(L_2) **For** each column number j **from** $c_{k+1} - 1$ **by** (-1) to $c_k + 1$ **do**
  (ℓ_2) **For** i **from** $l_k + 1$ **to** n **do**
    (a) $E := E_{i,j-1}(-\alpha_{ij}^* x^h)$; // here $a_{ij} = \alpha_{ij}^* x^{-r+h} + O(x^{-r+h+1})$ with $\alpha_{ij} \in$ F
    (b) $A := E[A]$; // This eliminates the term $\alpha_{ij}^* x^{-r+h}$ in a_{ij}
    (c) $T := TE$;
(3) **return** T;

Lemma 3.1. *Computing T and $T[A]$ (up to order ν) by the above algorithm costs* $O(\nu n^3)$.

Proof. Computing $E[A]$, in loop (ℓ_1), up to order ν requires $2n\nu$ multiplications in F. The cost of loop (ℓ_1) is equal to $2n\nu(n - c_k)$ and the cost of loop (L_1) is $2n\nu(n - c_k)(n_k - 1)$. Similarly, the cost of loop (L_2) is $2n\nu(n - l_k)(n_k - 1)$. The total cost is

$$2n\nu \sum_{k=1}^{d} (2n - c_k - l_k)(n_k - 1) \leq 2n\nu \sum_{k=1}^{d} (2n-1)(n_k - 1)$$
$$\leq 2n(2n-1)\nu \left(\sum_{k=1}^{d} n_k - d \right) = 2n(2n-1)(n-s-d)\nu \leq 4\nu n^3. \quad \square$$

We state the useful

Lemma 3.2. *Lifting the Arnold–Wasow form from order h to $h+1$ does not modify terms in A_L of valuation $\leq h+1$.*

Proof. Increasing the order in the Arnold–Wasow form is done using elementary transformations of the form $E_{ij}(\alpha_{ij}^* x^{h+1})$ with $\alpha_{ij}^* \in$ F as explained above. These elementary transformations induce row and column operations on $x^r A$ and A_L. Due to the factor x^{h+1}, terms of valuation $< h+1$ in A_L are not modified by these elementary transformations.
Terms in A_L of valuation $h+1$ can only be modified as a result of elementary transformations using rows or columns in $x^r A$ having valuation 0. The only non-zero constant terms in $x^r A$ are the 1s in the off-diagonal in A_0 (being in Jordan normal form) but the row and column positions of these 1s are, by definition, different from those positions of the elements in $x^r A$ that are used to form A_L. Hence elementary operations on $x^r A$, involving rows or columns of valuation 0 will not affect terms of valuation $h+1$ in A_L. This proves the lemma. $\square$

3.2. A normalised Moser-irreducible form

Lemma 3.3. *Suppose that A is Moser-irreducible. There exists a constant transformation C such that $\tilde{A} = C[A]$ with $\tilde{A}_0 = A_0$ and*

$$L(\tilde{A}, \lambda) = \begin{pmatrix} L^{11} & L^{12} & 0 \\ L^{21} & L^{22} - \lambda & 0 \\ L^{31} & L^{32} & J - \lambda \end{pmatrix} \tag{12}$$

where J is a nilpotent square matrix in Jordan normal form, L^{11} and L^{22} are square matrices of dimension d and $s-q$ with $0 \leq q \leq s$, and the matrix

$$\begin{pmatrix} L^{11} & L^{12} \\ L^{21} & L^{22} \end{pmatrix} \tag{13}$$

is nonsingular.

Proof. Let $q = 0$ and $L(A, 0)$ be partitioned

$$L(A,0) = \begin{pmatrix} L^{11} & L^{12} \\ L^{21} & L^{22} \end{pmatrix},$$

so that L^{11} and L^{22} are square matrices of order d and s respectively. If this matrix is nonsingular then put $C := I_n$ and we are done. Otherwise, the existence of the constant transformation C achieving (12) with $q > 0$ follows directly by reviewing the process of column-elimination algorithm as in the proof of Lemma 2.1, using the nonsingularity of the matrix in (13) as termination criterion, combined with a constant transformation of the form

$$\mathrm{diag}(I_{n-q}, S)$$

where $S \in \mathrm{F}^{q\times q}$ such that $S^{-1}L^{33}S = J$ (L^{33} as in (8)) is in Jordan normal form. □

Lemma 3.4. *Suppose that A is Moser-irreducible. Computing a constant transformation C such that $L(C[A], \lambda)$ has the structure as in Lemma 3.3, together with the ν first terms of $C[A]$ costs $\mathrm{O}(\nu n^3)$ multiplications in* F.

Proof. On the one hand, the computation of a constant transformation C' achieving (8) with the matrix (13) nonsingular together with the ν first terms of $A' := C'[A]$ requires $\mathrm{O}(\nu n^3)$ multiplications in F (see the proof of Lemma 2.4). On the other hand, L_{33} being a $q \times q$ nilpotent matrix, the computation of a constant matrix S such that $S^{-1}L^{33}S$ is in Jordan normal form costs $\mathrm{O}(q^3)$ operations in F (Storjohann, 2000; Storjohann and Villard, 2000). Finally computing $C''[A']$ (where $C'' := \mathrm{diag}(I_{n-q}, S)$) up to order ν costs $2\nu nq^2$ multiplications in F. Thus the total cost is of order $\mathrm{O}(\nu n^3)$ since $q \le s < n$. □

Definition 3.1. We say that a system (1) is in *normalised Moser-irreducible form* if it is Moser-irreducible with an L-matrix $L(A, \lambda)$ of the form as in Lemma 3.3.

A system that is in normalised Moser-irreducible form will be the prerequisite for the block-reduction algorithm as discussed in what follows.

In order to illustrate the notion of normalised Moser-irreducible forms, we present the following example: take the matrix

$$A = \begin{bmatrix} x^{-1} & x^{-3} & 2\,x^{-2} \\ x^{-2} & 0 & x^{-1}+x^{-2} \\ -x^{-2} & 0 & x^{-1}-x^{-2} \end{bmatrix}$$

where $n = 3 = r$. We have

$$A_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ -1 & 0 & -1 \end{bmatrix}.$$

The leading matrix A_0 is in Jordan normal form with $d = 1$, $n_1 = 2$ and $s = 1$.

The matrices A_L and $L(A, \lambda)$ are respectively:

$$A_L = \begin{bmatrix} x & x^2+x \\ -x & x^2-x \end{bmatrix}, \quad L(A,\lambda) = \begin{bmatrix} 1 & 1 \\ -1 & -1-\lambda \end{bmatrix}.$$

One has $\det L(A, \lambda) = -\lambda$, so the system $\vartheta y = Ay$ is Moser-irreducible. However it is not in normalised Moser-irreducible form since the matrix $L(A, 0)$ is singular. If we take $v = (-1, 1) \in \ker L(A, 0)$ (following the proof of Lemma 3.3) and consider the constant matrix

$$P = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

then we get

$$\tilde{A} = P[A] = \begin{bmatrix} -x^{-2} + x^{-1} & x^{-3} & 2\,x^{-2} \\ x^{-2} & 0 & x^{-1} \\ -x^{-2} & 0 & x^{-1} \end{bmatrix}$$

which is in normalised Moser-irreducible form with $q = 1$.

3.3. A block-reduced form

Theorem 3.1. *Consider a system (1) with Poincaré rank $r > 0$. Suppose that A is in normalised Moser-irreducible form with $L(A, \lambda)$ of the form (12) with $q \geq 1$. Then there exists an algorithm computing a transformation $T \in GL(n, \mathcal{O})$ with $T_0 = I$ such that $\tilde{A} := T[A]$ has the block-reduced (or block-triangular) form*

$$\tilde{A} = \begin{pmatrix} \tilde{A}^{11} & 0 \\ \tilde{A}^{21} & \tilde{A}^{22} \end{pmatrix} \tag{14}$$

with $\tilde{A}^{22}$ a square matrix of dimension q and $\tilde{A}_0 = A_0$, $\tilde{A}_{L,1} = A_{L,1}$.

Furthermore, the coefficients of T can be computed directly from the coefficients of A.

Remark 3.2. This algorithm can be easily extended to compute a block-diagonal form

$$\tilde{A} = \begin{pmatrix} \tilde{A}^{11} & 0 \\ 0 & \tilde{A}^{22} \end{pmatrix}$$

and gives hence an alternative method for proving the Generalised Splitting Lemma in the formulation as Lemma 4.1 in Pflügel (2004).

In what follows, we will say that the matrix A is block-reduced up to order $h \geq 0$ if

$$x^{-r}(A_0 + A_1 x + \cdots + A_h x^h)$$

is block-reduced as in (14). Similarly, for $h > 0$ the matrix A_L will be called block-reduced up to order h if

$$A_L = \begin{pmatrix} A_L^{11} & O(x^{h+1}) \\ A_L^{21} & A_L^{22} \end{pmatrix}$$

with A_L^{22} a square matrix of dimension q.

Before we give the proof of this theorem, note the following corollary which follows directly from Lemma 3.2:

Corollary 3.1. *Assume A is in Arnold–Wasow form up to order $h - 1$ ($h > 0$) and A_L is block-reduced up to order h. Then, lifting the Arnold–Wasow form to order h using elementary operations preserves the order h in the block-reduction of A_L.*

Proof (*Theorem* 3.1). We proceed by induction and assume that A is already block-reduced up to order $h - 1$, that is

$$x^r A = \begin{pmatrix} A^{11} & O(x^h) \\ A^{21} & A^{22} \end{pmatrix} \tag{15}$$

with the notation as in the theorem and $h > 0$, and that furthermore A_L is block reduced up to order h. For $h = 1$, these assumptions are satisfied if the system is in normalised Moser-irreducible form, so let $h > 1$.

In order to prove the theorem, we proceed in two steps. The fact that $x^r A$ is block reduced as in (15) implies that the upper-right block of A is in Arnold–Wasow form up to order $h - 1$. The first step consists in increasing the order of this partial Arnold–Wasow form to h whilst preserving the order h in the block-reduction of A_L. This follows directly from Corollary 3.1, using elementary operations.

The second step is to increase the order in the block-reduction of A_L to order $h+1$ without modifying the order h in the (partial) Arnold–Wasow form of A. Let

$$A_L = \begin{pmatrix} N^{11} & N^{12} \\ N^{21} & N^{22} \end{pmatrix}$$

with N^{11} a square matrix of dimension $d+s-q$ and N^{22} of dimension q. We have

$$N^{11} = N_1^{11} x + O(x^2) \tag{16}$$

with

$$N_1^{11} = \begin{pmatrix} L^{11} & L^{12} \\ L^{21} & L^{22} \end{pmatrix}$$

as in (13) nonsingular since A is in normalised Moser-irreducible form. Furthermore,

$$N^{12} = N_{h+1}^{12} x^{h+1} + O(x^{h+2}).$$

We show how we can eliminate the different columns of the leading matrix N_{h+1}^{12} using elementary transformations in the following order: assume we have eliminated the last $\tilde{q}$ columns ($0 \le \tilde{q} < q$) and we address the column of position $n-\tilde{q}$, denoted by $w \in \mathrm{F}^{d+s-q}$. Due to N_1^{11} being nonsingular, the columns of N_1^{11} form a basis of the vector space F^{d+s-q} and we can write w as a linear combination of these columns. Let $\sigma_i \in \mathrm{F}$ ($i = 1, \ldots, s+d-q$) be the coefficients of this linear combination. It is clear that the column operations that are induced by the sequence of elementary transformations $E_{c_i, n-\tilde{q}}(-\sigma_i x^h)$ eliminate w as desired. Let $R_{n-\tilde{q}}$ be the row in A of position $n-\tilde{q}$. The row operations induced by the elementary transformations add $R_{n-\tilde{q}}$, multiplied by $\sigma_i x^h$, to rows in A. Since the last row contains terms of order at least $O(x)$, these operations do not modify the order h in the partial Arnold–Wasow form of A. More specifically, the last $\tilde{q}$ coefficients of $R_{n-\tilde{q}}$ are of order at least $O(x^2)$ due to the structure of the matrix J in the normalised Moser-irreducible form being nilpotent in Jordan normal form. This makes sure that only columns in N_{h+1}^{12} of index $< n-\tilde{q}$ are altered and that the previous elimination of w is preserved. Finally, the element in $x^r A$ of position (c_i, n) is modified by adding $h\sigma_i x^{h+r}$, which does not modify the order h either since $r > 0$. We then increment $\tilde{q}$ by 1 and proceed with the remaining columns of N_{h+1}^{12}, until $\tilde{q} = q$. □

Remark 3.3. From this proof follows immediately an algorithm for the direct block-reduction. We draw attention to the fact that in the first step (Step 1 in the algorithm below), it is not necessary to carry out a lifting step of the Arnold–Wasow algorithm on the entire matrix A but only on the matrix A^{12}.

The algorithm can then be given as follows:

Direct_Block_Reduction(A, h)

Input: $A \in \mathcal{M}_n(\mathrm{K})$ in normalised Moser-irreducible form with parameter $q > 0$ and $h \in \mathbb{N}^*$

Output: $T \in GL(n, \mathcal{O})$ such that $B = T[A]$ is block-reduced up to order h (and furthermore, B_L is block-reduced up to order $h+1$)

(1) $T := I_n$;
(2) **For** k **from** 1 **to** h **do**
// A is block-reduced up to order $k-1$ and A_L is block-reduced up to order k
   (a) // Step 1: Increase the order of the partial Arnold–Wasow form in A^{12} to k
   Let E_1 be the transformation that corresponds to the sequence of elementary operations in the Arnold–Wasow algorithm achieving the increase in order;
   (b) $A := E_1[A]$; // A is now blocked-reduced up to order k
   (c) $T := TE_1$;

// Step 2: Increase the order of block-reduction in A_L to $k+1$

(a) **For each** column w_j of N^{12}_{k+1}, from column index $j = n$ **downto** $n - q + 1$ **do**
  (i) Compute its components σ_i, $(i = 1, \ldots, s + d - q)$, with respect to the basis of F^{d+s-q} formed by the columns of N^{11}_1;
  (ii) Apply successively the sequence of elementary transformations $E_{c_i,j}(-\sigma_i x^k)$;
(b) Update T accordingly;

(3) **return** T;

Lemma 3.5. *Suppose that A is in normalised Moser-irreducible form with $q \geq 1$. Let $h \geq 1$ be an integer. Then computing the ν first terms of a block-reduced form up to order h of A can be performed by using* $O(\nu h n^3)$ *multiplications in* F.

Proof. Following the proof of Theorem 3.1 and using the same notations, a block-reduced form up to order h can be obtained by executing the algorithm Direct_Block_Reduction with input parameters A and h.

By Lemma 3.1, Step 1 can be performed by using $O(\nu n^3)$ operations in F. In Step 2, task (i) can be performed by using Gauss elimination on the columns of the matrix $(N^{11}_1\ w_j)$ and hence costs $O((d + s - q)^3)$ (Von zur Gathen and Gerhard, 2003). In task (ii), we have to perform $d + s - q$ elementary transformations. Each elementary transformations requires $O(n\nu)$ multiplications in F, so the cost of this task is $O((d + s - q)n\nu)$. Finally, this is iterated q times in the loop in task (a) leading to an overall cost $O(q(d + s - q)n\nu)$, which is bounded by $O(\nu n^3)$. □

In order to illustrate the algorithm, we give an example. Consider the matrix

$$A = \begin{bmatrix} 0 & x^{-3} & 2x^{-2} \\ x^{-2} & 0 & x^{-1} \\ -x^{-2} & 0 & x^{-1} \end{bmatrix}$$

which is in normalised Moser-irreducible form with the associated matrix

$$A_L = \begin{bmatrix} x & x^2 \\ -x & x^2 \end{bmatrix}.$$

The matrix $x^3 A$ is block-reduced up to order 0, and A_L is block-reduced up to order 1. The elementary operation T_1 lifts the block-reduction to order 1, by lifting the order of the partial Arnold–Wasow form to 1, with

$$T_1 = E_{2,3}(-2x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2x \\ 0 & 0 & 1 \end{bmatrix}$$

and the transformed system

$$B := T_1[A] = \begin{bmatrix} 0 & x^{-3} & 0 \\ x^{-2} - 2x^{-1} & 0 & x^{-1} + 2 + 2x \\ -x^{-2} & 0 & x^{-1} \end{bmatrix}.$$

The elementary operation T_2 increases the order of the block-reduction in B_L to 2:

$$T_2 = E_{1,3}(-x) = \begin{bmatrix} 1 & 0 & -x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The resulting system is

$$C := T_2[B] = \begin{bmatrix} -x^{-1} & x^{-3} & x+2 \\ x^{-2} - 2x^{-1} & 0 & 4+2x \\ -x^{-2} & 0 & 2x^{-1} \end{bmatrix}.$$

Our implementation in ISOLDE handles this block-reduction algorithm in the following way: After entering the matrix A, we convert it to a power series matrix using

```
> B:=mat_convert(A,x,0);
```

We call the algorithm by providing B and some additional information such as r and q:

```
> C :=mat_direct_block_reduce(B,x,[4,1],1,1);
```

We now evaluate the matrix C up to order -1. Note that this corresponds to the lower left block A^{22} of the block-reduction process, which in this example is a matrix of dimension $n = q = 1$.

```
> mat_eval(C,x,-4,-1);
```

$$\left[4x^{-1} + 2x^{-2}\right]$$

From a theoretical point of view, this algorithm proves that the transformation matrix that achieves a block-reduction can be computed directly from the coefficients of the matrix A, without having to convert the system into a k-simple system as in Section 3.3 of Pflügel (2000). From a practical point of view, the algorithm based on the Generalised Splitting Lemma needs to calculate two transformation matrices S and T by solving *generalised Sylvester equations*. This is less efficient than using only one transformation T, based on elementary operations. The block-reduction algorithm can be implemented efficiently, using lazy evaluation of power series.

4. The super-reduction

4.1. A reduction criterion

A natural generalisation of the concept of Moser-reduction is to take into account additional coefficients of the matrix A. For $1 \le k \le r$ we define

$$m_k(A) = \max\left(0, r + \frac{\nu_0(A)}{n} + \frac{\nu_1(A)}{n^2} + \cdots + \frac{\nu_{k-1}(A)}{n^k}\right) \tag{17}$$

where ν_i denotes the number of columns of A having valuation $i - r - 1$ and

$$\mu_k(A) = \min_{T \in \mathcal{T}_n} \left\{ m_k\left(T[A]\right) \right\}.$$

Following Hilali/Wazner, we define:

Definition 4.1. The system (1) (the matrix A respectively) is called *k-reducible* if $m_k(A) > \mu_k(A)$, otherwise it is said to be *k-irreducible*. If the system is k-irreducible for $k = 1, \dots, r$, it is called *super-irreducible*.

It is well-known that by defining

$$\theta_k(A, \lambda) = x^{p_k(A)} \det\left(x^{r+1-k} A - \lambda I\right)\Big|_{x=0} \tag{18}$$

where

$$p_k(A) = k\nu_0(A) + (k-1)\nu_1(A) + \cdots + \nu_{k-1}(A).$$

one obtains a criterion to test whether a given system is k-reducible:

Proposition 4.1 (*Hilali and Wazner, 1987*). *For all $1 \le k \le r$, the system (1) is k-irreducible if and only if the polynomials $\theta_j(A, \lambda)$, $(j = 1, \dots, k)$, do not vanish identically in λ.*

Proposition 4.2. *Assume that the system is Moser-irreducible and block-reduced as in (14). Then the matrix A is super-irreducible if and only if the matrix A^{22} is.*

Proof. This follows from the fact that A^{11} is already super-irreducible and that for a block-reduced matrix we have

$$\theta_k(A, \lambda) = \theta_k(A^{11}, \lambda) \cdot \theta_k(A^{22}, \lambda).$$

Hence A is super-irreducible if and only if $\theta_k(A^{22}, \lambda)$ does not vanish. □

4.2. The algorithm

We now describe our new algorithm to compute a super-irreducible form of a given system of the form (1) with size n and Poincaré-rank r. The algorithm consists in iterating the following:

Super_reduction(A)
  Input: $A \in \mathcal{M}_n(K((x)))$
  Output: $T \in \mathcal{T}_n$ such that $T[A]$ is super-irreducible

(1) **if** $n = 1$ **or** $r = 0$ **or** A_0 is nonsingular **then return**(I_n);
(2) **if** A_0 is not nilpotent **then**
  (a) Apply the *Block-diagonalisation* algorithm in Barkatou (1997) (see the appendix) and assign T to the resulting transformation matrix;
  (b) Perform recursion on the resulting two blocks A^{11} and A^{22} and denote by T^{11} and T^{22} the resulting transformation matrices
  (c) **return**($T \cdot \mathrm{diag}(T^{11}, T^{22})$);
(3) **else if** A is not Moser-irreducible **then**
  // One can now assume that the matrix A_0 is nilpotent;
  (a) Apply our Moser-reduction algorithm to get an equivalent normalised Moser-irreducible matrix and denote by T_1 the resulting transformation matrix;
  (b) **return**($T_1 \cdot T_2$) where T_2 is obtained by recursively calling the algorithm on the Moser-irreducible system;
(4) **else** // The system is in normalised Moser-irreducible form
  (a) Let q be the parameter of the normalised Moser-irreducible form
  (b) **if** $q = 0$ **then return**(I_n);
  (c) **else**
    (i) Apply the direct block-reduction algorithm to compute an equivalent block-reduced form[3] as in Theorem 3.1 and denote by T the resulting transformation matrix;
    (ii) Perform recursion on the second block $\tilde{A}^{22}$ and denote by T^{22} the resulting transformation matrix;
    (iii) **return**($T \cdot \mathrm{diag}(I_{n-q}, T^{22})$);

Note that the matrix $\tilde{A}^{22}$ is of Poincaré-rank $< r$ and size $q < n$. Iterating this process we can compute, in at most $\min(n-1, r)$ steps, an equivalent super-irreducible matrix which is upper block-triangular

$$\tilde{A} = \begin{pmatrix} \tilde{A}^{1,1} & & 0 \\ \vdots & \ddots & \\ \tilde{A}^{\tilde{r},1} & \dots & \tilde{A}^{\tilde{r},\tilde{r}} \end{pmatrix} \qquad (19)$$

where the individual diagonal blocks are super-irreducible.

4.3. Complexity analysis

We are now able to give the complexity of our super-reduction algorithm.

Proposition 4.3. *Consider a system* (1) *with size n and Poincaré-rank r. Then the above algorithm computes an equivalent super-irreducible form up to order ν using at most* $O(\nu r n^4 \min(n-1, r))$ *multiplications in* F.

---

[3] Up to a big enough order h. It is sufficient to take $h = \nu = nr$.

Comparison of methods A and B				
	Poincaré-rank r			
n	2	3	4	5
21	9 15	26 26	67 208	232 253
22	32 12	33 30	65 119	198 384
23	13 13	42 46	184 99	674 732
24	28 12	35 19	124 218	317 1079
25	34 14	94 38	96 98	951 1045
26	20 15	51 55	180 166	350 655
27	35 19	52 26	312 270	1496 330
28	38 23	93 48	300 185	1136 1043
29	63 47	73 75	422 333	1460 2515

Fig. 1. Comparison of Method A and Method B.

Proof. By Theorem A.2, step 2 can be performed by using $O(\nu n^3)$ operations in F. By Proposition 2.2, step 4 costs $O(\nu r n^4)$. Finally, step 6 costs $O(\nu h n^3)$ which is bounded by $O(\nu r n^4)$ if we take $h \leq nr$. Now the process has to be repeated at most $\min(n-1, r)$ times, so the total cost is $O(\nu r n^4 \min(n-1, r))$. □

4.4. An example

In order to demonstrate our implementation, we re-visit the previous example.

```
> B:=mat_convert(A,x,0);
```

We call the super-reduction with B as input:

```
> super_reduce(B,x,u);
```

$$[A23, [[1, u-2], [2, -u], [3, u^3]], [4, 1]]$$

The output is a list of the calculated super-irreducible matrix, and local invariants such as the integer slopes of the Newton-polygon of B and the associated Newton-polynomials (see Pflügel (2000) for a more detailed discussion of this).

5. Discussion

Our algorithm has been implemented in the current version of ISOLDE (Barkatou and Pflügel, 2006). Fig. 1 shows the results of a performance comparison between two different methods for the super-reduction: We have tested matrix dimensions for n from 21 to 29. For each dimension, a series of test matrices having Poincaré-rank $r = 2$ to 5 has been run with our method from Barkatou and Pflügel (2007) (Method A, first entry in each cell) and the method from this paper (Method B, second cell entry). The times are given in seconds, on a PC with 2 GHz and 1 GB RAM using Maple 11 on Windows 2000.

From this data, we draw the following conclusions:

- For increasing values of n, Method B (the direct block-reduction) seems to be faster than Method A (generalised block-reduction algorithm, using the Generalised Splitting Lemma) except for some exceptional cases.
- The greater the Poincaré-rank r, the larger one needs to choose n for Method B to be faster.

- The presented data makes it somewhat difficult to spot a clear trend, but we strongly suspect that a review of the data structures used for the lazy evaluation mechanism by our implementation would increase the performance and reduce the number of occurring outliers.

The Moser- and Super-reduction being a fundamental building block of the algorithms computing formal solutions (or any kind of global solutions which require local information), our algorithm will be beneficial for a whole range of symbolic algorithms for systems of linear differential equations.

Acknowledgements

We would like to thank Gary Broughton for his assistance with the implementation of our algorithms in Maple and the anonymous referees for their helpful comments.

Appendix. Two splitting lemmas and their complexity

In this section, we give the complexity of two algorithms that are used by our algorithm **Super_reduction**. The complexity analysis is straightforward but does not seem to have been published elsewhere.

A.1. Classical splitting lemma

Consider an $n \times n$ matrix A with entries in K

$$A(x) = x^{-r}(A_0 + A_1 x + A_2 x^2 \cdots)$$

with $A_0 \neq 0$ and $r > 0$.

In this section we assume that A_0 is not nilpotent and we explain how to compute a transformation matrix $T \in \mathrm{GL}(n, \mathcal{O})$ which transforms the matrix A into an equivalent block-diagonal matrix $B = x^{-r}\mathrm{diag}(B^{11}, B^{22})$ such that B_0^{11} is non singular and B_0^{22} nilpotent.

Lemma A.1. *By a constant similarity transformation C the leading matrix A_0 can be changed into a block-diagonal matrix*

$$C^{-1}A_0C = B_0 = \mathrm{diag}(B_0^{11}, B_0^{22})$$

with B_0^{11} nonsingular and B_0^{22} nilpotent in Jordan form. This can be done in $\mathrm{O}(\nu\, n^3)$ operations in F.

Proof. This can be done in the following way:

(1) Compute the matrix ${A_0}^n$ (costs $\mathrm{O}(n^\omega \log n)$ using binary powering which is bounded by $\mathrm{O}(n^3)$).
(2) Compute a basis $\{P_1, \ldots, P_d\}$ of the space generated by the columns of ${A_0}^n$ (costs $\mathrm{O}(n^3)$ using Gauss elimination).
(3) Compute a basis $\{P_{d+1}, \ldots, P_n\}$ of the kernel of ${A_0}^n$ (costs $\mathrm{O}(n^3)$ using Gauss elimination).
(4) Form the matrix P whose columns are $P_1, \ldots, P_n$ and compute its inverse P^{-1} (costs $\mathrm{O}(n^3)$).
(5) Compute $P^{-1}A_0P = \mathrm{diag}(\tilde{A}_0^{11}, \tilde{A}_0^{22})$ with $\tilde{A}_0^{11}$ nonsingular and $\tilde{A}_0^{22}$ nilpotent (costs $\mathrm{O}(n^3)$).
(6) Compute Q such that $Q^{-1}\tilde{A}_0^{22}Q$ be in Jordan form (costs $\mathrm{O}(q^3)$, where q denotes the size of $\tilde{A}_0^{22}$ (Storjohann, 2000).
(7) Compute $C = P\ \mathrm{diag}(I_p, Q)$ and $C^{-1} = \mathrm{diag}(I_p, Q^{-1})P^{-1}$ (costs $\mathrm{O}(2pq^2 + q^\omega)$ which is bounded by $\mathrm{O}(n^3)$).
(8) Compute $C^{-1}AC$ up to order ν, (costs $\mathrm{O}(\nu\, n^\omega)$). □

Theorem A.1. *Consider a matrix*

$$A(x) = x^{-r}(A_0 + A_1 x + \cdots)$$

in $\mathcal{M}_n(\mathrm{K})$. *Assume that*

$$A_0 = \mathrm{diag}(A_0^{11}, A_0^{22})$$

such that A_0^{11} *and* A_0^{22} *have no eigenvalues in common. Then there exists a matrix* $T(x) = I + T_1 x + T_2 x^2 + \cdots$ *in* $\mathrm{GL}(n, \mathrm{K})$ *such that*

$$T[A] = x^{-r}(A_0 + B_1 x + \cdots + B_i x^i + \cdots)$$

and each B_i *is block-diagonal matching the partition of the matrix* A_0. *Furthermore for all* $i \in \mathbb{N}$, *the matrices* T_{i+1} *and* B_{i+1} *only depend on* $A_0, A_1, \ldots, A_i$.

A constructive proof of this result is given in Wasow (1967) (cf. Section 12, pp. 52–54). For completeness we will give the idea of the proof here:

Proof. We put $T_0 = I$ and $B_0 = A_0$ and look for matrices T_i of the special form

$$T_i = \begin{pmatrix} 0 & T_i^{12} \\ T_i^{21} & 0 \end{pmatrix}.$$

Then one can show that for $i \geq 1$ the coefficients T_i and B_i can be obtained by successively solving $2i$ equations of the form $MX - XN = U$ or $NY - YM = V$ where $M = A_0^{11}$, $N = A_0^{22}$ and where U and V depend only on A_j, B_j, T_j for $j = 0, \ldots, i-1$. □

The above results lead to a block-diagonalisation algorithm. It takes as input a matrix $A(x) = x^{-r}(A_0 + A_1 x + \cdots) \in \mathcal{M}_n(\mathrm{K})$ and an integer ν and returns a matrix $T(x) = T_0 + T_1 x + \cdots + T_\nu x^\nu$ with $\det T_0 \neq 0$ and a block-diagonal matrix $B(x) = x^{-r}(B_0 + B_1 x + \cdots + B_\nu x^\nu)$ with $B_0 = \mathrm{diag}(B_0^{11}, B_0^{22})$, B_0^{11} non singular and B_0^{22} nilpotent, such that $T[A] - B = O(x^{-r+\nu+1})$.

This proof uses the following well-known result: Let M and N be two square matrices of order p and q with entries in the field F whose characteristic polynomials are relatively prime (i.e. having no common eigenvalues in the algebraic closure $\overline{\mathrm{F}}$ of F) then for any $p \times q$ matrix in F, there exists a unique $p \times q$ matrix X in F such that $MX - XN = U$.

Lemma A.2. *Let* M *and* N *be two square matrices of order* p *and* q *with entries in* F. *Suppose that* M *is nonsingular and* N *is nilpotent in Jordan form*

$$N = \begin{pmatrix} 0 & \varepsilon_1 & & 0 \\ 0 & \ddots & & \\ & & \ddots & \varepsilon_q \\ 0 & & 0 & 0 \end{pmatrix}, \quad \varepsilon_i \in \{0, 1\}.$$

Let U *be a* $p \times q$ *matrix with entries in* F. *Then the solution* X *of the Sylvester equation* $MX - XN = U$ *can be computed in* $\mathrm{O}(n^3)$ *operations in* F *where* $n = \max(p, q)$.

Proof. The Sylvester equation $MX - XN = U$ admits a unique solution X which can be determined by solving the following q linear systems:

$$MX_1 = U_1, \quad \text{and} \quad MX_j = \varepsilon_{j-1} X_{j-1} + U_j \quad \text{for} \quad j = 2, \ldots, q$$

where X_j (*resp.* U_j) denotes the jth column of X (U *resp.*). This follows from the fact that the columns of XN are successively $0, \varepsilon_1 X_1, \ldots, \varepsilon_{q-1} X_{q-1}$.

To solve the above systems, it suffices to compute the inverse M^{-1} of M and then perform successively the operations :

$$X_1 = M^{-1} U_1, \quad X_j = M^{-1}\left(\varepsilon_{j-1} X_{j-1} + U_j\right), \quad \text{for} \quad j = 2, \ldots, q.$$

The inverse of M can be computed in $\mathrm{O}(p^\omega)$. Each column X_j can be computed in $\mathrm{O}(p^2)$ operations in F. So X can be computed in $\mathrm{O}(qp^2) + \mathrm{O}(p^\omega)$ operations in F which is bound by $\mathrm{O}(n^3)$. □

Theorem A.2. *With the notations above, the cost of the block-diagonalisation algorithm as in Theorem A.1 is* $\mathrm{O}(\nu n^3)$ *operations in* F.

Proof. By a constant similarity transformation C, the leading matrix A_0 can be changed into a block-diagonal matrix

$$C^{-1}A_0C = \mathrm{diag}(A_0^{11}, A_0^{22})$$

with A_0^{11} nonsingular and A_0^{22} nilpotent in Jordan form. This requires $\mathrm{O}(\nu n^3)$ operations in F.

In order to compute T and B, one has to solve 2ν Sylvester equations of the form $MX - XN = U$ or $NY - YM = V$ where $M = A_0^{11}$ and $N = A_0^{22}$. Since N is nilpotent in Jordan form then we need at most $\mathrm{O}(\nu n^3)$ operations in F. □

A.2. Generalised splitting lemma

We assume now that A_0 is nilpotent and A is Moser-irreducible with associated polynomial θ. The Generalised Splitting Lemma computes a transformation matrix that yields a block-diagonal matrix $B = x^{-r}\mathrm{diag}(B^{11}, B^{22})$ such that B_0^{11} is nilpotent and $B^{22} = B_1^{22}x + O(x^2)$ and $\det(B_1^{22} - \lambda I) = \theta(\lambda)$.

Theorem A.3. *Consider a system as in* (1)*, let* θ *be its associated polynomial and put* $d = \deg(\theta)$*. Then the system is equivalent to a system*

$$\vartheta z = x^{-r}\begin{pmatrix} B^{11} & 0 \\ 0 & B^{22} \end{pmatrix} z$$

where $B^{11} \in \mathcal{M}_{n-d}(\mathcal{O})$ *and* $B^{22} \in \mathcal{M}_d(\mathcal{O})$*. Furthermore,* B^{11} *is Moser-irreducible with constant associated polynomial,* B_0^{11} *is nilpotent and* $\mathrm{rank}\,(B_0^{11}) = \mathrm{rank}\,(A_0)$*, and* $B^{22} = B_1^{22}x + O(x^2)$*.*

Proof. A proof of this theorem can be found in Pflügel (2004), using the fact that a system that is Moser-irreducible can be rewritten as a so-called k-simple system of the form

$$D(x)\vartheta z = x^{-k}N(x)z \tag{A.1}$$

where $k = r - 1$ and $D, N \in \mathcal{M}_{n-d}(\mathcal{O})$ with $\det(N_0 - \lambda D_0) \neq 0$. The proof is constructive, it essentially generalises the proof of the classical Splitting Lemma (Theorem A.1) to k-simple systems, using the concept of regular matrix pencils. It requires solving a generalised type of Sylvester equations using the following result: Let $M - \lambda P$ and $N - \lambda Q$ be two regular matrix pencils of order p and q with entries in F. Then the matrix equation

$$(M - \lambda P)X - Y(N - \lambda Q) = R - \lambda S$$

has a unique solution $X, Y \in F^{p\times q}$ for any given $R, S \in F^{p\times q}$ provided the two regular matrix pencils $M - \lambda P$ and $N - \lambda Q$ have no common eigenvalue. A proof of this result can be found in Chu (1987). □

The above results lead to a generalised block-diagonalisation algorithm. It takes as input a matrix $A(x) = x^{-r}(A_0 + A_1x + \cdots) \in \mathcal{M}_n(\mathrm{K})$ and an integer ν and returns a matrix $T(x) = T_0 + T_1x + \cdots + T_\nu x^\nu$ with $\det T_0 \neq 0$ and a block-diagonal matrix $B(x) = x^{-r}(B_0 + B_1x + \cdots + B_\nu x^\nu)$ with $B = \mathrm{diag}(B^{11}, B^{22})$, B_0^{11} nilpotent and B^{22} a matrix of dimension $n - q$ and Poincaré-rank $r - 1$ such that $T[A] - B = O(x^{-r+\nu+1})$.

Theorem A.4. *Using the above notations, the cost of the generalised block-diagonalisation algorithm (as in the proof of Theorem A.3) is* $\mathrm{O}(\nu n^3)$ *operations in* F.

Proof. In order to apply the Generalised Splitting Lemma, the leading pencil $N_0 - \lambda D_0$ of (A.1) needs to be brought into a block-diagonal form (see Pflügel (2000)). It is well known that the computation of normal forms of regular matrix pencils can be reduced to classical matrix normal forms (see e.g. Quéré-Stuchlik (1996)), so the cost will be $O(n^3)$ operations in F.

In order to compute T and B, one has to solve 2ν generalised Sylvester equations of the form

$$(M - \lambda P)X - Y(N - \lambda Q) = R - \lambda S.$$

Solving these equations can be reduced to solving a linear system of dimension $2pq$ which can be done in $\mathrm{O}(n^3)$ operations in F (here $p, q < n$). Overall, the transformations need to be computed for the first ν coefficients of the system similarly as in the classical Splitting Lemma case which gives a total cost of $\mathrm{O}(\nu \cdot n^3)$. □

References

Babitt, D., Varadarajan, V., 1983. Formal reduction of meromorphic differential equations: A group theoretic view. Pacific J. Math. 109 (1), 1–80.

Barkatou, M., 1993. An algorithm for computing a companion block diagonal form for a system of linear differential equations. Journal of App. Alg. in Eng. Comm. and Comp. 4, 185–195.

Barkatou, M., 1997. An algorithm to compute the exponential part of a formal fundamental matrix solution of a linear differential system. Journal of App. Alg. in Eng. Comm. and Comp. 8 (1), 1–23.

Barkatou, M., 1999. On rational solutions of systems of linear differential equations. J. Symbolic. Comput. 28, 547–567.

Barkatou, M., Pflügel, E., 2006. The ISOLDE package. A SourceForge Open Source project, http://isolde.sourceforge.net.

Barkatou, M., Pflügel, E., 2007. Computing super-irreducible forms of systems of linear differential equations via Moser-reduction: A new approach. In: Proceedings of ISSAC'07. ACM Press, Waterloo, Canada, pp. 1–8.

Chu, K., 1987. The solution of the matrix equations AXB - CXD = E and (YA - DZ,YC - BZ) = (E,F). Linear Algebra Appl. 93, 93–105.

Churchill, R.C., Kovacic, J.J., 2002. Cyclic vectors. In: Rutgers University, N. (Ed.), In Differential Algebra and Related Topics, Proceedings of the International Workshop. World Scientific Publishing Co., River Edge, NJ.

Cluzeau, T., 2003. Factorization of differential systems in characteristic *p*. In: Proceedings of ISSAC '03. ACM, New York, NY, USA, pp. 58–65.

Dietrich, V., 1978. Zur Reduktion von linearen Differentialgleichungssystemen. Math. Ann. 237, 79–95.

Giesbrecht, M., Storjohann, A., 2002. Computing rational forms of integer matrices. J. Symbolic. Comput. 34 (3), 157–172.

Hilali, A., Wazner, A., 1987. Formes super-irréductibles des systèmes différentiels linéaires. Numer. Math. 50, 429–449.

Moser, J., 1960. The order of a singularity in Fuchs' theory. Math. Z. 379–398.

Pflügel, E., 1998. Résolution symbolique des systèmes différentiels linéaires. Ph.D. Thesis, LMC-IMAG.

Pflügel, E., 2000. Effective formal reduction of linear differential systems. Appl. Alg. Eng. Comm. Comp. 10 (2), 153–187.

Pflügel, E., 2004. An improved estimate for the maximal growth order of solutions of linear differential systems. Arch. Math. 83, 256–263.

Quéré-Stuchlik, M., 1996. Théorie des faisceaux linéaires de matrices, Liens avec la théorie des systèmes linéaires et Application à la résolution d'équations algébro-différentielles. Ph.D. Thesis, LMC–IMAG.

Storjohann, A., 2000. Algorithms for matrix canonical forms. Ph.D. Thesis, Swiss Federal Institute Of Technology.

Storjohann, A., Villard, G., 2000. Algorithms for Similarity Transforms. Extended Abstract. Bregenz, Austria, pp. 109–118.

Von zur Gathen, J., Gerhard, J., 2003. Modern Computer Algebra. Cambridge University Press, New York, NY, USA.

Wasow, W., 1967. Asymptotic Expansions for Ordinary Differential Equations. Robert E. Krieger Publishing.